

Role of Clinical Trials Informatics in the NCI's Cancer Informatics Infrastructure

John Silva, MD, and Robert Wittes, MD,
National Cancer Institute, Bethesda, Maryland

Abstract

A Web-enabled Cancer Informatics Infrastructure (CII) will enable faster, better clinical trials, ultimately improving cancer care. To create the CII, the National Cancer Institute is forming public-private partnerships and building on existing activities. Key innovations include development of standards for cancer patient information and clinical research, management of patients using disease specific "clinical states" and use of drag and drop electronic protocol authoring.

Introduction

Today cancer research faces many problems. Patients and physicians must overcome tremendous paper and administrative "hassle factors." Clinical trials produce eclectic data focused on the specific trial rather than a patient-centered record, most of which is paper-based and handled manually. There is no systems approach to moving discoveries into routine care, and there is no national-scale infrastructure to support trials or basic research.

Recent advances in information technology and the explosive growth of the Web can overcome these problems in health care and specifically in cancer research. With growing number of users connected to the Internet, businesses are using electronic commerce and financial services to stay competitive in the "zero time" market place. Web-based information technologies have brought fundamental changes to how corporate America works. Learning from these models, the National Cancer Institute proposes to create a Cancer Informatics Infrastructure (CII) that will exploit the most promising discoveries from its research base, speed clinical trials and make the best possible cancer care routine for all Americans.

Concept

The goals of the CII are fourfold. First, it will create an enterprise for both basic and clinical research in cancer, effectively integrating research at the molecular level into research at the patient's side. Second, it will develop usable standards with the private sector to ensure interoperability. Third, it will simplify clinical trials, supporting protocol development and significantly shortening the time

from discovery to trial. Fourth, it will provide secure information tailored to the user needing it, effectively meeting the requirements of trialists, patients, providers, insurers, employers, and other authorized users.

For example, by using their computer, patients and their oncologists can find, for the patient's specific cancer, the best treatments and clinical trials. The physician can fill in a simple form, submit it, and - if the physician and patient meet certain standards and criteria - the patient will be entered into the clinical trial, receive the therapy, and remain in the care of his or her physician. This will remove many of the barriers patients and physicians face in entering clinical trials. Slow and cumbersome paper-based systems of collecting data for multi-center studies will give way to electronic communication, linking sites of care delivery (hospitals, doctors' offices and clinics) and secure, research databases of investigator's. At the same time, by accruing patients to trials faster than before, answers to key scientific questions will come much faster.

It is the specific intent of the CII to collect patient related data only once, in a secure fashion. Information required to select the best care available or to determine eligibility for a specific clinical trial will be sent from the physician's local electronic patient record system. Information required specifically by the clinical trial (for example, case report forms) will be automatically generated from the local systems used to document care delivered to the patient. This is planned to eliminate the overwhelming redundancies of today's "paper hassles and chases".

To achieve these goals, the clinical trials enterprise will exploit existing and emerging technologies, while observing certain principles. *User-focused*, it will address the needs of multiple constituencies. *Simple*, it will appear seamless to its users. *Private*, it will protect the patient's rights to privacy and confidentiality and investigator intellectual property. *Standard*, it will use common terminology to ensure interoperability. The principle of *value* is paramount- the enterprise system will provide immediate, tangible value to its users.

Initial CII activities are focused on treatment trials: surgical and radiation therapy trials for local disease, adjuvant therapy trials for regional disease and trials for established metastatic disease. (Figure 1, shaded area). In subsequent stages, NCI will incorporate the entire cancer spectrum into the CII. The result will be revolutionary. For the first time, all cancer clinical trials will have common terminology and reporting requirements, greatly increasing the speed and efficiency of conducting a trial and the accuracy of its results.

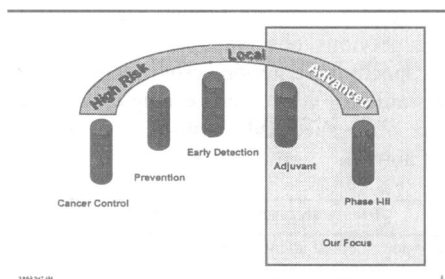


Figure 1. The cancer spectrum

Today, paper-based case report forms (CRFs) pass through multiple checks and reviews and are shipped back and forth by Federal Express. As the first step toward easier clinical trials, the CII proposes to replace this time-consuming paper chase with remote data entry using a secure Web connection. (Figure 2)

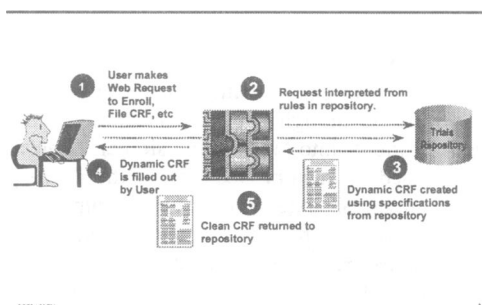


Figure 2. Remote data entry

The request to enroll and file a CRF will be sent to the clinical trials repository. The repository will interpret the request and generate a dynamic CRF for the user to fill out. Corrections will be completed at the time of filing, according to specifications in the repository, and the "clean" CRF will be returned to the repository. Studies indicate that the dynamic

CRF, in addition to being quicker than the existing paper-based system, will also have the benefit of decreasing error rates. (Figure 3)

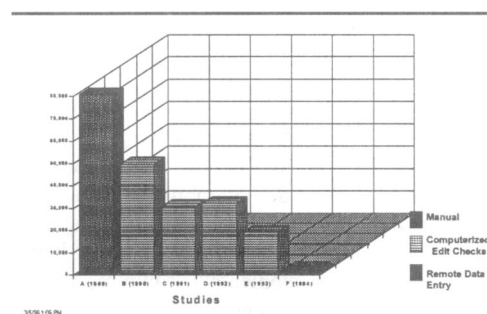


Figure 3. Case report form error rates

Next steps will include capturing clinical trial data during patient-practitioner encounters, then generating clean CRFs from clinical data repositories. Hassle factors will be eliminated as payors and administrators accept (and pay for) health service requests which were specified in pre-approved clinical trial protocols. (Figure 4)

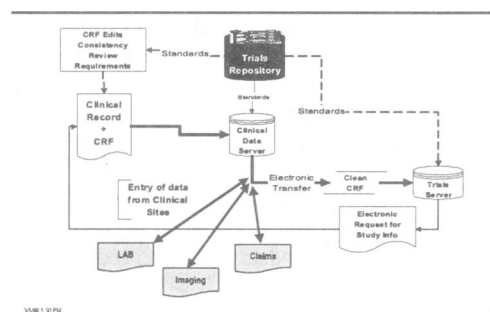


Figure 4. Framework for tomorrow's easier trials

Approach

To realize this ambitious vision, the NCI is forming public-private partnerships to develop a national-scale blueprint for the CII architecture. This involves benchmarking successes in other markets and convincing vendors and others in the private sector that participating in the development and deployment of the CII will indeed benefit them. These efforts are especially important for two reasons. First, NCI expects to contribute the minority of the total funding

that the fully realized CII will require. Second, the private sector will be where almost all of the work of building the CII will take place. NCI also plans to take advantage of information planning tools to create standardized clinical trials “Lego Blocks” and promote development the informatics specifications and reference architecture for the CII.

NCI Activities

Within the NCI, work in progress is laying the groundwork for the CII. NCI enterprise systems are being established to provide common models across the institute and to ensure that data and applications are sharable. To simplify CRFs, project teams have identified common data elements for clinical trials. Automated document management is addressing bottlenecks in the paper-based system, and efforts are underway to develop concurrent protocol authoring and to use drag and drop authoring for electronic protocol generation that will enable users to spend most of their time on the front end, thinking about scientific questions and how the trial is going to answer them.

Clinical Trials Models

Initial work at NCI analyzed of over 80 active clinical trials across multiple cancers to identify commonalities and develop a generic “clinical trial model”. Six major categories of information were derived: eligibility, efficacy, schema of the trial, safety, informed consent and the scientific question that the trial proposes to address. (Figure 5) Detailed entity relationship diagrams and an object oriented model were crafted from the data elements specified (or implied) by the trial document and CRFs. Common data elements were constructed from the set of all data elements in an attempt to harmonize, standardize and simplify data collection requirements [1].

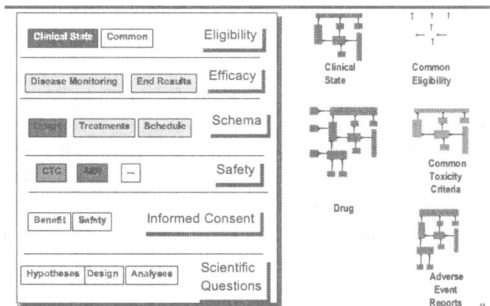


Figure 5. The generic clinical trial

Eligibility Analysis

As part of initial attempts to develop standards for eligibility for clinical trials, inclusion and exclusion criteria from twenty three active phase II and III breast cancer trials (11 adjuvant and 12 metastatic disease) were evaluated. Over 520 distinct and, occasionally conflicting, requirements were identified and classified into a multi-level eligibility hierarchy whose top-level nodes were: patient characteristics, primary tumor characteristics, nodal status (extent of regional disease), metastatic status and prior therapy. Figure 6 depicts some of the various ways trial authors describe the presence of adenocarcinoma of the breast (under primary tumor characteristics). Confirming previous observations, we documented variance in specifying criterion across trials that were otherwise similar in intent. For example, laboratory values were often specified as absolute or as some multiple of normal.

<i>Adenocarcinoma of the Breast</i>
Histologically documented adenocarcinoma of the breast
Histologically confirmed adenocarcinoma of the breast
Histologically documented diagnosis of mammary adenocarcinoma with evidence of recurrent or metastatic disease
Histologically documented mammary adenocarcinoma
Diagnosed primary invasive adenocarcinoma of the breast
Histologically confirmed diagnosis of breast cancer
Operable, invasive primary breast cancer
Diagnosed primary invasive adenocarcinoma of breast

Figure 6: Eligibility criteria for primary tumor characteristics

Clinical States

To assist eligibility determinations, we have established the concept of a clinical state. A clinical state can best be thought of as a mechanism for segregating patients into “relatively” similar, clinical groupings based on history, features of the primary tumor and regional extent, metastatic status and response to prior therapy. The states are chosen to divide the clinical trials and diagnostic and treatment protocols such that these will be similar for all patients in that state. The mechanics and representation of clinical states for prostate cancer is described in a companion paper [2]. A national panel of clinicians, pathologists, trialists, advocates and statisticians created the clinical state model for breast cancer depicted below. (Figure 7)

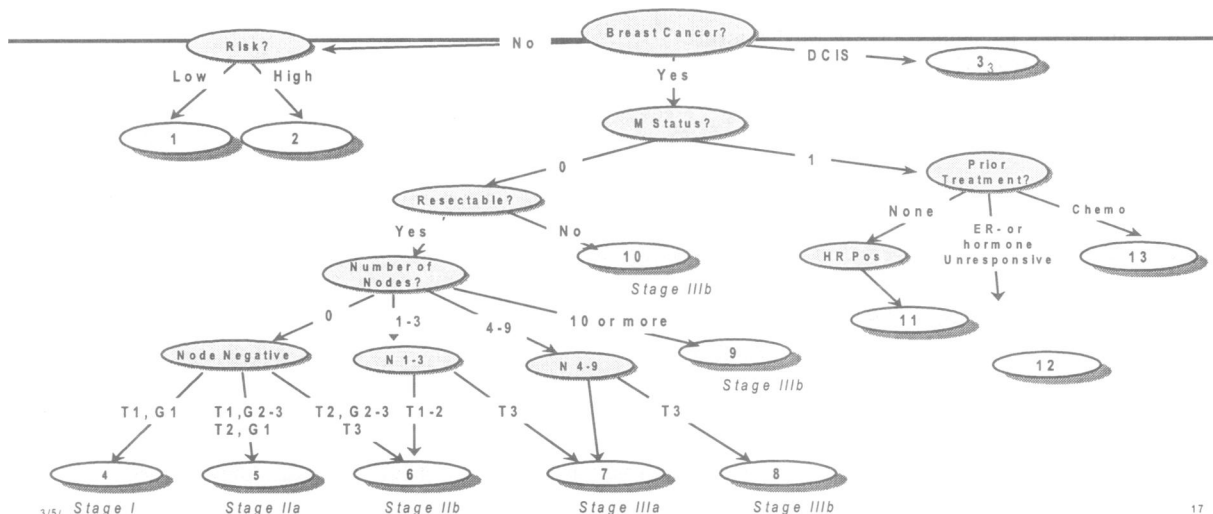


Figure 7. Breast cancer clinical states

clinical trials, such as the recently described GuideLine Interchange Format [2].

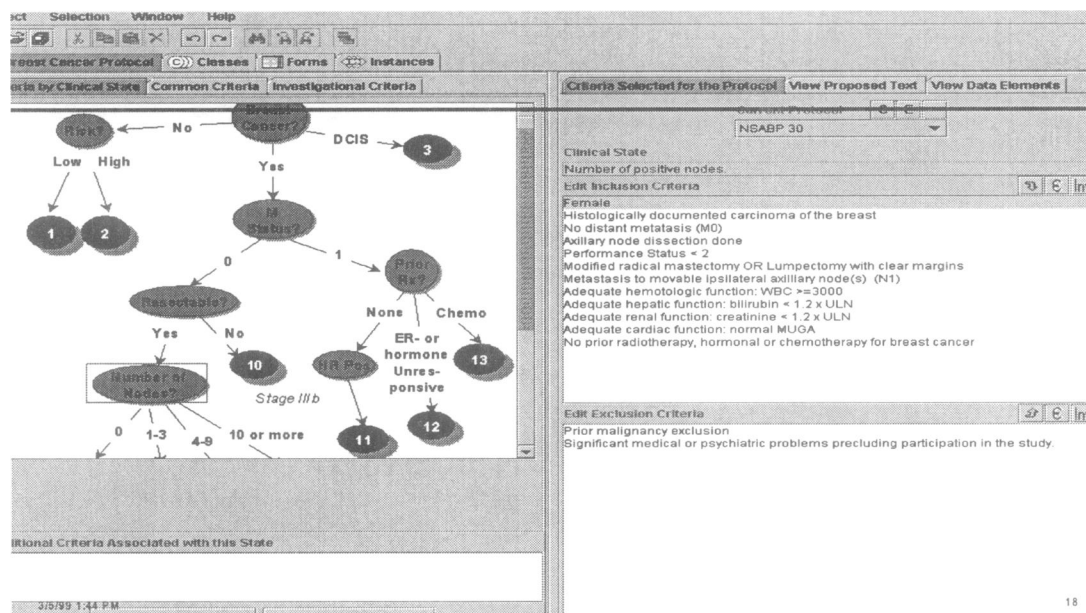


Figure 8 Protégé representation of breast cancer clinical states.

Figure 8 demonstrates the core set of eligibility/exclusion criteria for one clinical state in breast cancer. Each criterion is linked to its corresponding common data element. We are currently working with Cooperative groups, national experts and patient advocates to standardize both the common data element set and their definition. NCI will continue to elaborate common data elements for all cancers and to explore methods for representing and sharing

References

1. Common data element dictionary available at <http://hiip-wrkstn.hpc.org>. Follow links in Projects to Common Data Elements
2. Rubin, DL, Gennari, JH, Srinivas, S. et al. Tool Support for Authoring Eligibility Criteria for Cancer Trials. Submitted to AMIA 1999.
3. Ohno-Machado, L, Gennari, JH, Murphy, SN, et. al. The Guideline Interchange format: A model

for representing guidelines. J Am Med Inform Assoc. 1998; 5:357-372.